DOCUMENT RESUME

ED 317 606 TM 014 715

AUTHOR Clark, Sheldon B.
TITLE The Noggin Factor in Survey Research: Developing New Techniques for Assessing Nonresponse Bias.
INSTITUTION Oak Ridge Associated Universities, Tenn.
SPONS AGENCY Department of Energy, Oak Ridge, TN. Technical Information Center.; National Science Foundation, Washington, D.C.
PUB DATE Apr 90
CONTRACT DE-AC05-76OR00033; SRS-8809065
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Engineers; Immigrants; National Surveys; *Research Problems; Scientists; Statistical Bias; *Statistical Surveys
IDENTIFIERS National Science Foundation; *Noggin Factor; *Nonrespondents; Survey Research

ABSTRACT

The primary objective of this paper is to encourage survey researchers not to become overly reliant on the literature for generic solutions to non-response bias problems. In addition, the paper recounts an example of how a non-traditional approach was used to maximize the usefulness of data collected under unusual constraints and with an a priori expectation of a high rate of non-response. The author was charged with testing the ability of the National Science Foundation to conduct a biennial survey of scientists and engineers who had recently immigrated into the United States. This assessment was done within the context of the "noggin factor," which engenders a consideration of the particularity or each survey situation. As a result of the noggin factor, generalized solutions to survey research problems tend to be self-defeating. It is concluded that survey researchers should attend more to process than to outcome. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED317606

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

SHELDON B. CLARK

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

The Noggin Factor in Survey Research: Developing New Techniques for Assessing Nonresponse Bias

by

Sheldon B. Clark
Oak Ridge Associated Universities

Paper Presented at the Annual Meeting of
the American Educational Research Association
Boston, Massachusetts
April 16-20, 1990

TM014715

BEST COPY AVAILABLE

The Noggin Factor in Survey Research: Developing New Techniques for Assessing Nonresponse Bias[1]

Introduction

It is not uncommon for survey researchers working in applied settings to ignore the issue of potential bias associated with nonresponse when they relate the findings of their efforts to the decision-makers or clients to whom they are accountable or to others interested in their research. How many researchers have had drafts of reports returned by technical editors or clients with all of their caveats, qualifiers, and meticulous delineations of the limitations of a study "red-penned" or criticized as being "too academic?" Presumably nonresponse bias is a technical problem that these "bottom line" consumers of this research cannot relate to.

By the same token, how often are authors of scholarly papers using survey research data apparently guilty of the same attitude? One might be tempted to lump the practice of not reporting considerations of nonresponse bias with that of not reporting experimentwise error rates and to dismiss both as being minutia that are not relevant to the objectives of the study. The researcher who calculates the experimentwise error rate for all the statistical tests he/she performs in the course of an investigation (i.e., not just those he/she chooses to include in the final paper!), may be horrified at the result. Likewise the survey researcher who analyzes questionnaire data based on a 30 percent, 75 percent, or even 95 percent response rate may be in for a rude awakening when he/she uses a nonstochastic (i.e., worst possible case) model to compute confidence intervals for the observed item means (see Clark & Nichols, 1983, or Cochrane, 1977, p. 361).

Many mathematical and nonmathematical procedures for estimating the potential effects of nonresponse bias have been suggested, tested, and compared (see Berdie, Anderson, & Niebuhr, 1986; Donald, 1960; Kalton, 1983; Madow, Nisselson, & Olkin, 1983; Madow & Olkin, 1983; Madow, Olkin, & Rubin, 1983). For many researchers, however, the mathematical procedures are too complex, too conservative, or too esoteric to be useful, and the validity of the nonmathematical approaches may be suspect. Unfortunately such judgments often result in the failure to address this very important question at all.

It should be noted that nonresponse bias is a problem **only when bias exists**--that is, when the respondents differ from nonrespondents **on some characteristic that is relevant to the purposes of the survey**. Berdie (1989) contends that "an obsessive fear of nonresponse bias is not justified."

---

[1]This material is based upon activities supported by the National Science Foundation under Agreement No. SRS-8809065 with the U.S. Department of Energy. Oak Ridge Associated Universities operates under the U.S. Department of Energy Contract No. DE-AC05-76OR00033. Any opinion, findings and conclusions, or recommendations expressed in this paper are those of the author and do not necessarily reflect the policies and views of the National Science Foundation, the U.S. Department of Energy, or Oak Ridge Associated Universities.

Such a view is only reasonable, however, when the researcher is familiar enough with the target population to make informed judgments (e.g., based on previous experience with the same or a similar population) about the characteristics of the nonrespondents and the relationship between those characteristics and the survey topic. In many applied settings it is not uncommon to have just such experience to draw upon.

It has also been suggested that researchers may be overlooking readily available means of making at least a cursory examination of the existence and likely effects of nonresponse bias. For example, Clark and Finn (1989) recommend that investigators use any information they have about all the individuals surveyed to gain insight into how the respondents compare to the nonrespondents on selected variables for which they have information. Depending on what information is available and how it relates to the known and unknown characteristics of the two groups relative to the domain covered in the questionnaire, such an analysis may or may not offer any clues as to what sorts of nonsampling bias exist.

Suggestions like the above are consistent with a philosophy of conducting survey research that Babbie supports: the means by which one can become a more competent and more versatile survey researcher is not to collect a bigger stack of survey research "cookbooks," but to focus on developing the logic and skills of survey research. "No survey satisfies the theoretical ideals of scientific inquiry. Every survey represents a collection of compromises between the ideal and the possible. . . . *Perfect* surveys may not be possible, but *good* surveys can and should be done." (p. iii)

This paper has two objectives: (a) to encourage survey researchers not to become overly reliant on the literature for generic solutions to problems and opportunities that are in fact unique to their particular applications, and (b) to recount an example of how a nontraditional approach was used to maximize the usefulness of data collected under unusual constraints and with the a priori expectation of a high rate of nonresponse.

The Noggin Factor

The term "survey research" is a very broad one which encompasses mail surveys, telephone surveys, face-to-face interviews, and all sorts of variants and combinations of these techniques. Even within one of these categories, the content, purposes, methods, and target populations vary so much from one application to another that it seems almost absurd to try to make generalizations about the comparative effectiveness of different approaches. Anyone who has tried to search the literature for research on particular aspects of questionnaire design, sampling, and other survey-related topics is keenly aware of the diversity of academic fields in which this research might appear and the often inconsistent and contradictory findings reported in studies within and across these fields.

Many explanations have been offered for these apparent discrepancies in the literature, including basic differences in the studies with regard to such factors as the nature of the survey population, the subject matter of the questionnaire, and the quality of the survey instrument itself. The contradictory findings of ostensibly comparable studies have led some authors to attempt to establish general guidelines for questionnaire development and even rating systems by which the quality of questionnaires can be assessed (Boser & Clark, 1990; Clark & Boser, 1989). Not surprisingly, such efforts have had only limited success. The characteristics found to be desirable for all mail questionnaires were mostly of the "motherhood-and-apple-pie" variety.

This author believes that generalized solutions to survey research problems tend to be self-defeating and that more attention should be given to the unique characteristics of the survey research application at hand--not only the characteristics of the target population, availability of resources, and so forth, but also the more basic question of what the researcher hopes to accomplish by conducting the survey. The primary concern of survey research efforts is not always to estimate population parameters as precisely as possible, so it is not reasonable to operate as if that were the case. Since resources are limited in most survey applications, compromises and trade-offs have to be made in such areas as sample size; number and type of follow-ups; content, format, and appearance of questionnaires; duplication process; and mailing mode. Researchers should be more conscientious in considering what they hope to accomplish with a particular effort when they make these trade-offs and compromises. They should be more creative in developing solutions to their problems by taking advantage of whatever special opportunities particular applications may offer. These considerations are what we refer to as "the noggin factor."

There are few, if any, universal truths in survey research. This is not to say that paying attention to good survey research cannot enhance the skills of survey researchers in settings that are different from those in particular studies. On the contrary, the noggin factor implies that those researchers who are familiar with the widest variety of survey research have the most "experience" on which to draw. The noggin factor does suggest, however, that researchers might make greater strides in improving their craft by viewing published survey research in a different light. Rather than searching for systems that might be applied to their own research, they should perhaps pay more attention to each component of those systems.

The Noggin Factor: A Case Study

The author was charged with testing the feasibility of the National Science Foundation's (NSF's) carrying out a biennial survey of scientists and engineers who had recently immigrated to the United States. Information about the need for the survey and how the proposed survey fits into NSF's mission, as well as technical details about the survey's development, implementation, results, and

conclusions are beyond the scope of the present paper.[2] For the purpose at hand, we shall consider only those elements of the effort that help illustrate the approach previously termed "the noggin factor."

Many of the characteristics of this task were unique, and some were especially problematic. For example:

- There is very little reliable empirical data on surveys of new immigrants and virtually none on the targeted subset of that population.

- The only reliable data source for the names and addresses of new immigrants is maintained by the U.S. Immigration and Naturalization Service (INS).

- In recent years over 600,000 individuals annually become permanent residents, of which less than 5 percent are thought to be scientists and engineers.

- The occupational data (on which, conceptually, the determination of whether an individual is part of the target population should be made) in the INS base was thought to be unreliable by those most familiar with it. This uncertainty made it advisable to survey a population of which more than 95 percent would be highly unlikely to relate to the content of the questionnaire--a situation which by itself would likely lead to very low response rates.

- In addition to the problem of the expected lack of interest in the survey topic for a large portion of the new immigrants, it was known that many of them would be unable to respond to the questionnaire due to educational, cultural, and most importantly, language barriers.

- Because the names and addresses of new immigrants are protected by the Privacy Act of 1974, it was not possible for the author's organization to participate directly in the sample selection or the initial mailing. Since there was no way to identify individuals included in the sample drawn, it was not possible to have normal follow-ups of nonrespondents. (INS could have undertaken the follow-up effort, but factors unrelated to the project prevented them from doing so.)

Operating under constraints such as these renders a <u>systems</u> approach to survey research design essentially worthless. Given the limitations plus the omnipresent time and money restrictions, the most important question to be answered was, ***"What do we hope to accomplish in this endeavor?"*** After careful deliberation the research priorities emerged. Among them were the following:

---

[2]The reader who is interested in these aspects of the project are referred to Clark and Finn (1990).

- To determine the reliability of the INS occupational data. This determination would allow subsequent stratified sampling strategies to be carried out with more confidence.

- To gather information which would be helpful in redefining the target population in the future. (The possibilities for redefinition centered around certain characteristics of the individuals measured by INS--e.g., class of admission, nonimmigrant class of entry, occupational category.)

- To gain insight into how well this population responds to government surveys and how this response varies by demographic characteristics. This information would be valuable (a) in determining the feasibility of carrying out a reliable survey of this population (or some subset of it) at all, and (b) in allocating resources in any future survey endeavor.

By establishing these priorities, it became clear that our basic concern was not with parameter estimation. In effect, our research objectives were primarily exploratory--we wanted to find out as much as possible about a broad range of new immigrants in order to establish an empirical basis for designing subsequent surveys. Without dwelling on every factor and alternative considered, suffice it to say that we eventually devised an approach which we felt was consistent with our priorities (given the project's constraints), and which was likely to result in the fewest rival hypotheses. Decisions were based on the best information available. This information came from a variety of sources and was of varying quality: empirical data from our own research experience and from the research experiences of others; anecdotal information from our own experience and from that of others; "educated guesses" based on theoretical and hypothesized relationships; intuition--our own and that of others; and bare-faced guesses. Several attributes of this approach are relevant to the discussion at hand.

For NSF's purposes, the primary characteristic of interest in the INS data base was occupation. Even though the reliability of this data element on the INS data base was questionable, it was plausible that certain INS occupational categories had a greater probability of including NSF-defined scientists and engineers than others. On the basis of extensive experience with occupational data and special analyses of existing INS data, we divided the 29 INS occupational categories into 3 groups, from which to draw the sample of 6,000.[3] The highest sampling rate (66 percent; n=2,500) was applied to the group most likely to be scientists or engineers; the next highest sampling rate (10 percent; n=500) was applied to the group likely to contain a much lower proportion of scientists and engineers; and the lowest sampling rate (3 percent; n=3,000) was applied to the very large remaining group unlikely to contain a significant proportion of scientists and engineers. While the relationship of these sampling rates may at first seem counterintuitive, it was consistent with our stated priorities.

---

[3]For detailed information about the sampling frame, the reader is referred to Clark and Finn (1990).

Given our research priorities and prior knowledge of the target population, the overall rate of response was less important than getting responses from the complete spectrum of subgroups of interest. Of course, in order to have reliable information about these subgroups, it was necessary to get as many responses as possible despite the many negative factors operating in some subgroups. Among the techniques used to maximize response rate, two are relevant to the present discussion.

The sampling frame consisted of those foreign nationals who recently applied for and had been granted permanent residence status by INS. For such immigrants the tangible proof of this action is the "green card," which has significant positive employment and reporting implications and is a valued and necessary means of identification. It is reasonable to assume that the INS envelope that contained the "green card" was a harbinger of good news for most of these individuals. To take advantage of this recent positive interaction with INS, the initial contact for this survey was made in an envelope with an INS return address, and the introductory letter from NSF referred to INS's support of, and interest in, the survey effort.

Another technique used to elicit information from as many potential respondents as possible was to first solicit minimal information in an easy-to-respond format. This was particularly important in this study, because there was no mechanism for identifying individuals included in the initial mailing. Once an individual had revealed his/her identity through an initial simple inquiry, we could include him/her in traditional monitoring and follow-up procedures once the full questionnaire had been sent. Included in the initial mailing was a brief cover letter, which explained the purpose of the survey and INS's role in it, and a postage-paid postcard on which the recipient was asked to check one of the following boxes:

- ❑ By education and/or current employment, I am a natural or social scientist or engineer, and I am willing to participate in the mail survey.
- ❑ By education and/or current employment, I am a natural or social scientist or engineer, but I do not wish to participate in the mail survey.
- ❑ By education and/or current employment, I am not a natural or social scientist or engineer, but I am willing to participate in the mail survey.
- ❑ By education and/or current employment, I am not a natural or social scientist or engineer, and I do not wish to participate in the mail survey.

The card also had the individual's name and address affixed, and he/she was asked to make any necessary corrections.

Besides the desire to obtain the names and addresses of as many recipients of the initial mailing as possible, the reasons for using this two-stage process (i.e., having a return postcard in one mailing and then sending the questionnaire in a second mailing) included the following:

- Even if a questionnaire were ultimately not returned, we would still have one potentially useful piece of information supplied by the individual (i.e., whether he/she perceives himself/herself as a scientist or engineer). We had no information about how reliable this self-assessment might be, and there was no basis for estimating the proportion of questionnaires we might expect from those who indicate a willingness to participate. By collecting this minimal information, however, we at least left ourselves with some options.

- The possibility of a respondent's indicating an unwillingness to participate in the survey was believed to have a positive effect on the response rate for postcards. One could argue, however, that it might have had a negative effect on the response rate for questionnaires, since some of those who declined to participate in the survey might actually have completed a questionnaire, had they not been given the opportunity to decline receiving one.

- There was some concern that the questionnaire itself could have a negative effect on response rate, because it might look complicated to the recipient, and because it was clearly aimed at those with science or engineering backgrounds. After all, we were knowingly sampling a large number of individuals who were unlikely to identify with the content and purpose of the survey.

In cooperation with INS, a mechanism was designed by which we could compare certain characteristics of the nonrespondents with those of the respondents without violating the terms of the Privacy Act. These comparisons allowed us to assess the likelihood of nonresponse bias, since information on the most likely source of bias in this context--occupation--was available for respondents and nonrespondents alike. An identification number on the address label allowed us to get aggregate information from INS about those who returned postcards versus those who did not.

For the most part, the results of this effort are not germane to the present discussion. In the interest of closure, however, a few relevant outcomes are noted:

- We found INS occupational data to be much more reliable than had been expected. Coupled with the patterns found between occupations and several immigrant classifications, this will allow the target population to be redefined to eliminate many of the "nonproductive" categories of immigrants. Such a redefinition will significantly decrease the size of the target population, making it more homogeneous and thereby increasing

the likelihood of achieving higher response rates. It will also enable researchers who conduct future surveys of this population for similar purposes to allocate their resources more judiciously.

- Almost all of the conjectures about how these individuals might respond to this type of survey were supported by the data. For example, (a) natives of some countries were much more likely to respond than those of others; (b) educational and language barriers existed more for some subgroups of the population than for others; and (c) those who identified with the content and purpose of the survey were much more likely to respond: at all; positively to the invitation to participate in the survey; and to the questionnaire.

- The distribution of postcard responses across the four response options presented earlier was 53 percent, 7 percent, 19 percent, and 21 percent, respectively. The groups most likely to contain scientists and engineers were also the most likely to indicate a willingness to participate in the survey.

- Response rates were calculated for various identifiable subsets of the sample. These rates ranged from 10 percent to 90 percent, with the higher rates being associated with those groups having a high proportion of scientists and engineers, and the lower rates being associated with those groups having a very small proportion of scientists and engineers.

- There was no evidence that the nonrespondents differed from the respondents in any meaningful way for a survey aimed at collecting occupational data from new immigrants.

- Of the respondents who were scientists or engineers according to NSF's criteria, over 98 percent had so-indicated on the initial postcard. This meant that the self-assessments were very reliable, and some general conclusions could be drawn about the 22 percent of the postcard respondents who had indicated a willingness to participate in the survey but who ultimately did not return a questionnaire.

The focus of the present paper has been on the positive aspects of the use of "the noggin factor" in a survey research application, but, in the interest of intellectual honesty, it should be noted that not all outcomes of this study were positive and expected. As in most applied research, all viable alternatives were not identified during the design phase, unanticipated problems arose during the course of the study, and the results raised some questions that could not be answered with confidence.

Conclusion

Too often survey researchers slavishly choose their procedures from among those that they have found to be effective in the past or from those that are presented in the literature. It is hoped that this paper will encourage these

researchers to approach each survey effort more critically and creatively, seeking viable solutions to (and/or taking full advantage of) any circumstances that are unique to the application at hand. In addition, we hope that the ideas presented here will stimulate the community of survey researchers to view the survey efforts of others a little differently--not as being potential sources of "technique packages" for immediate or wholesale adoption in their own research, but as an idea pool from which they can draw particular elements in certain situations.

Perhaps we should pay more attention to the process rather than the outcomes of the survey research of others, since the constraints and objectives of our own applications are likely to differ from those of other applications in ways that may significantly affect the outcomes. We should evaluate the research of others in the context in which it was conducted--including the objectives and constraints that were in play. We should share information about how, and according to what criteria, available resources are allocated to the achievement of various objectives. We should discuss the knowns and unknowns, the experience base, and the logic that underlie our methods. By sharing these experiences with other researchers, the knowledge-base and, more importantly, the **idea-base** in survey research can be extended, and we can each enhance our own survey research skills.

References

Babbie, E.R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing Co.

Berdie, D.R. (1989). Reassessing the value of high response rates to mail surveys. Marketing Research, 1(3), pp. 52-64.

Berdie, D.R., Anderson, J.F., & Niebuhr, M.A. (1986). Questionnaires: design and use (2nd ed.). Metuchen, N.J.: Scarecrow Press, Inc.

Boser, J.A. & Clark, S.B. (1990, April). Consensus on desirable characteristics of mail questionnaires: illusion or reality? Paper presented at the annual meeting of the American Educational Research Association, Boston.

Clark, S.B., & Boser, J.A. (1989, March). Seeking consensus on empirical characteristics of effective mail questionnaires: A first step. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Document Reproduction Service No. ED 306 305)

Clark, S.B., & Finn, M.G. (1989, March). Interpreting nonresponse in survey research: methodological heresy? Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Document Reproduction Service No. ED 308 208)

Clark, S.B., & Finn, M.G. (1990). Improving immigration and emigration statistics on scientists and engineers (Oak Ridge Associated Universities Publication No. ORAU-90/A-95). Oak Ridge, TN: Oak Ridge Associated Universities.

Clark, S.B., & Nichols, J.O. (1983, May). Increasing the precision of estimates in follow-up surveys: a case study. Paper presented at the 23rd annual forum of the Association for Institutional Research, Toronto.

Cochrane, W.G. (1977). Sampling techniques (3rd ed.). New York: John Wiley & Sons.

Donald, M.N. (1960). Implications of nonresponse for the interpretation of mail questionnaire data. Public Opinion Quarterly, 24(1), pp. 99-114.

Fowler, F.J. (1984). Survey research methods. Beverly Hills: Sage.

Kalton, G. (1983). Compensating for missing survey data. Ann Arbor Michigan: Institute for Social Research.

Madow, W.G., Nisselson, H., & Olkin, I. (Eds.). (1983). Incomplete data in sample surveys: Vol. 1. Report and case studies. New York: Academic Press.

Madow, W.G., & Olkin, I. (Eds.). (1983). Incomplete data in sample surveys: Vol. 3. Proceedings of the symposium. New York: Academic Press.

Madow, W.G., Olkin, I., & Rubin, D.B. (Eds.). (1983). Incomplete data in sample surveys: Vol. 2. Theory and bibliographies. New York: Academic Press.